

Dynamic adaptive binning: an improved quantification technique for NMR spectroscopic data

Paul E. Anderson · Deirdre A. Mahle ·
Travis E. Doom · Nicholas V. Reo ·
Nicholas J. DelRaso · Michael L. Raymer

Received: 24 May 2010/Accepted: 6 September 2010
© Springer Science+Business Media, LLC 2010

Abstract The interpretation of nuclear magnetic resonance (NMR) experimental results for metabolomics studies requires intensive signal processing and multivariate data analysis techniques. A key step in this process is the quantification of spectral features, which is commonly accomplished by dividing an NMR spectrum into several hundred integral regions or bins. Binning attempts to minimize effects from variations in peak positions caused by sample pH, ionic strength, and composition, while reducing the dimensionality for multivariate statistical analyses. Herein we develop an improved novel spectral quantification technique, dynamic adaptive binning. With this technique, bin boundaries are determined by optimizing an objective function using a dynamic programming strategy. The objective function measures the quality of a bin configuration based on the number of peaks per bin. This technique shows a significant improvement over both traditional uniform binning and other adaptive binning techniques. This improvement is quantified via synthetic validation sets by analyzing an algorithm's ability to create

bins that do not contain more than a single peak and that maximize the distance from peak to bin boundary. The validation sets are developed by characterizing the salient distributions in experimental NMR spectroscopic data. Further, dynamic adaptive binning is applied to a ^1H NMR-based experiment to monitor rat urinary metabolites to empirically demonstrate improved spectral quantification.

Keywords NMR · Metabolomics · Binning · Quantification · Dynamic programming

1 Introduction

Metabolomics, the measurement of metabolite concentrations and fluxes in various biological systems, is one of the most comprehensive of all bionomics (Fiehn 2002; Reo 2002). Unlike proteomics and genomics that assess intermediate products, metabolomics assesses the end product of cellular function, metabolites. Changes occurring at the level of genes and proteins (assessed by genomics and proteomics) may or may not influence a variety of cellular functions. But metabolomics, by contrast, assesses the end products of cellular metabolic function, such that the measured metabolite profile reflects the cellular metabolic status. For instance, a disease or foreign compound may interfere at the genomic or proteomic level, while it will always manifest itself at the metabolomic level. Further, nuclear magnetic resonance (NMR) spectroscopy of biofluids has been shown to be an effective method in metabolomics to identify variations in biological states (Lindon et al. 2001; Nicholson and Wilson 1989; Shockcor and Holmes 2002). In contrast to various other proteomic, genomic, and metabolomic analyses, NMR spectroscopy is

Electronic supplementary material The online version of this article (doi:[10.1007/s11306-010-0242-7](https://doi.org/10.1007/s11306-010-0242-7)) contains supplementary material, which is available to authorized users.

P. E. Anderson · T. E. Doom · M. L. Raymer
Department of Computer Science and Engineering,
Wright State University, Dayton, OH 45435, USA

P. E. Anderson (✉) · D. A. Mahle · N. J. DelRaso
Air Force Research Laboratory, Biosciences and Protection
Division, Wright-Patterson AFB, Dayton, OH 45433, USA
e-mail: anderson.51@wright.edu

D. A. Mahle · N. V. Reo
Department of Biochemistry and Molecular Biology,
Wright State University, Boonshoft School of Medicine,
Cox Institute, Dayton, OH 45429, USA

Report Documentation Page			Form Approved OMB No. 0704-0188	
<p>Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p>				
1. REPORT DATE 25 NOV 2010	2. REPORT TYPE	3. DATES COVERED		
4. TITLE AND SUBTITLE Dynamic adaptive binning: an improved quantification technique for NMR spectroscopic data			5a. CONTRACT NUMBER	
			5b. GRANT NUMBER	
			5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)			5d. PROJECT NUMBER	
			5e. TASK NUMBER	
			5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Wright State University, Dayton, OH, 45435		8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSOR/MONITOR'S ACRONYM(S)	
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited.				
13. SUPPLEMENTARY NOTES				
14. ABSTRACT <p>The interpretation of nuclear magnetic resonance (NMR) experimental results for metabolomics studies requires intensive signal processing and multivariate data analysis techniques. A key step in this process is the quantification of spectral features, which is commonly accomplished by dividing an NMR spectrum into several hundred integral regions or bins. Binning attempts to minimize effects from variations in peak positions caused by sample pH, ionic strength, and composition, while reducing the dimensionality for multivariate statistical analyses. Herein we develop an improved novel spectral quantification technique, dynamic adaptive binning. With this technique, bin boundaries are determined by optimizing an objective function using a dynamic programming strategy. The objective function measures the quality of a bin configuration based on the number of peaks per bin. This technique shows a significant improvement over both traditional uniform binning and other adaptive binning techniques. This improvement is quantified via synthetic validation sets by analyzing an algorithm's ability to create bins that do not contain more than a single peak and that maximize the distance from peak to bin boundary. The validation sets are developed by characterizing the salient distributions in experimental NMR spectroscopic data. Further, dynamic adaptive binning is applied to a H NMR-based experiment to monitor rat urinary metabolites to empirically demonstrate improved spectral quantification.</p>				
15. SUBJECT TERMS				
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES 12
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified		19a. NAME OF RESPONSIBLE PERSON

non-invasive, non-destructive, and requires little sample preparation (Reo 2002).

Typically, NMR metabolic spectroscopic data are analyzed as follows: (1) standard post-instrumental processing of spectroscopic data, such as the Fourier transformation, phase adjustment, and baseline correction; (2) quantification of spectral features commonly implemented via binning; (3) normalization and scaling; and (4) multivariate statistical modeling of data. Quantification of spectral features, step (2), is a key step in the development of classification algorithms and biomarker identification (i.e., pattern recognition). A common method of quantification employed by the NMR community is known as binning or bucketing, which divides an NMR spectrum into several hundred regions. This technique is performed to (1) minimize effects from variations in peak positions caused by sample pH, ionic strength, and composition (Spraul et al. 1994); and (2) reduce the dimensionality for multivariate statistical analyses. The result is a data set with fewer features, thereby, increasing the tractability of pattern recognition techniques, such as principal component analysis (PCA) (Hotelling 1933; Jolliffe 1986) and partial least squares discriminant analysis (PLS-DA) (Martens and Naes 1989; Wold 1966).

There are several alternatives to spectral binning that still provide data dimension reduction. Examples of these include PARS (Forshed et al. 2005; Torgrip et al. 2003), curve-fitting method for direct quantification (Crockford et al. 2005), peak alignment tools in HiRes (Zhao et al. 2006), and targeted profiling (Weljie et al. 2006). These techniques identify peaks or specific peak patterns in the spectra that are conserved across spectra. After the patterns have been identified, they are quantified by determining the peak area or amplitude. The accuracy of these algorithms is dependent on the spectral resolution, the quality of the peak alignment, and the breadth of spectroscopic pattern databases. Since spectral resolution is dependent upon the magnetic field strength (i.e., instrument specific), the spectral patterns in complex mixtures (e.g., urine and plasma) are also field dependent. This adds another level of complexity to targeted profiling techniques that attempt to match spectral patterns against standard spectra acquired at a specific magnetic field.

Recently, there have been several full resolution techniques developed (Cloarec et al. 2005; Stoyanova et al. 2004) and applied (Schoonen et al. 2007, b). In most cases, these techniques require the spectra to be preprocessed by an alignment algorithm, providing a “cleaner” data set (Forshed et al. 2002, 2003, 2005; Stoyanova 2004; Vogels et al. 1993, 1996). Further, several examples of artifacts from unaligned NMR signals have been reported, thus, proper alignment is a critical problem when applying full resolution techniques (Brekke et al. 1989; Brown and

Stoyanova 1996; Defernez and Colquhoun 2003; Forshed et al. 2002; Stoyanova et al. 2004), and is necessary to correct for differences that reflect variations of the individual’s metabolism (Nicholson et al. 1999). In contrast, bin-based techniques attempt to mitigate misalignment by dividing the spectrum into regions that ideally remove quantification errors due to misalignment.

Despite the development of these differing quantification techniques, binning remains a common high throughput quantification technique for the NMR community (Åberg et al. 2009). Additionally, with the continually expanding applications of NMR-based metabolomics into new fields with variations of sample types, this generalized binning methodology continues to be an appropriate first approach. Thus new techniques that improve this methodology, such as that described herein, remain an important avenue for development. For a detailed comparison of alternative quantification techniques to binning see Åberg et al. 2009; Cloarec et al. 2005; Weljie et al. 2006.

The traditional binning method is to divide a spectrum into several hundred non-overlapping regions or bins of equal size. This simple technique has been shown to be effective in the field of metabolomics (Beckwith-Hall et al. 2002; Beckwith-Hall et al. 1998; Connor et al. 2007; Gartland et al. 1990; Griffin et al. 2001; Robertson et al. 2000; Robosky et al. 2002; Wang et al. 2004; Whitehead et al. 2005). While uniform binning mitigates the effects from variations in peak positions, shifts occurring near the boundaries can result in dramatic quantitative changes in the adjacent bins due to the non-overlapping boundaries. This problem can be countered by incorporating a kernel-based binning method that weights the contribution of peaks by their distance from the center of the bin (Anderson et al. 2008).

Another method for countering this problem is to dynamically determine the size and location of each bin. One such dynamic binning algorithm is adaptive intelligent binning, which recursively identifies bin edges in existing bins (De Meyer et al. 2008). Another dynamic binning method is adaptive binning, which uses the undecimated wavelet transform to smooth a composite spectrum. The observed peaks and minima of the smoothed composite spectrum are then used to dynamically bin the spectra (Davis et al. 2007). The composite spectrum is smoothed to remove multiple observed peaks that arise from misaligned peaks. When using a composite spectrum, adjacent peaks in the same spectrum can be misinterpreted as resulting from shifts from a single peak. To overcome this drawback, the bin boundaries can be dynamically determined by optimizing a heuristic based objective function that utilizes individual spectra smoothed via a wavelet transform.

Herein we propose a novel dynamic binning method, dynamic adaptive binning, for processing NMR spectroscopic data for multivariate analysis. With this technique

bin boundaries are dynamically determined via dynamic programming by optimizing an objective function that measures the quality of the bin configuration. This technique is shown to be superior to the traditional uniform binning technique and other advanced binning techniques (adaptive binning and adaptive intelligent binning) based on their ability to create bins containing a single peak and maximize the distance from peak to bin boundary. This comparison is facilitated by synthetic data sets that capture the salient characteristics of ^1H NMR spectroscopic data from a urinary profile (Anderson et al. 2009). Finally, a case study demonstrates the capabilities of dynamic adaptive binning in comparison to uniform binning on a ^1H NMR-based experiment to monitor rat urinary metabolites.

2 Methods

The technique of spectral binning is a general signal processing technique that reduces the dimensionality of spectroscopic data while attempting to retain the pertinent information and mitigate quantitative effects of peak misalignment. Spectral quantification transforms every sample, represented as an NMR spectrum, into a feature vector. Biomarker identification can then be defined as finding a set of features that describe a pattern between groups, thus, the success of biomarker identification is directly related to the quality of the feature vectors. Here a biomarker is defined as a set of NMR signals that change relative to some reference (i.e., before and after exposure to a toxin). Such an experiment will have at least two groups (e.g., pre-dose and post-dose) for which spectroscopic data is compiled.

2.1 Spectroscopic data

Both empirical and synthetic spectroscopic data are employed to show the application of dynamic adaptive binning. The synthetic spectroscopic data sets are based on urine ^1H spectra and were developed by characterizing the salient distributions in empirical spectroscopic data (Anderson et al. 2009). Each spectrum is modeled as a combination of Gaussian–Lorentzian peaks and a piecewise cubic interpolated baseline. These synthetic data sets enable the use of exacting performance metrics because the true location and size of each peak is known *a priori*. By using the synthetic data sets, metrics are developed that directly measure the ability of a spectral binning algorithm to create bins containing a single observed peak, while minimizing the probability of splitting peaks between bins. In addition to comparing spectral binning algorithms on synthetic data sets, this manuscript demonstrates the application of dynamic adaptive binning on empirical data

from a ^1H NMR-based experiment to monitor rat urinary metabolites after exposure to α -naphthylisothiocyanate (ANIT).

Animals were given a single administration, via oral gavage at 10 ml/kg, of ANIT in corn oil vehicle at one of the following doses: 20, 50, and 100 mg/kg. Control animals received corn oil only at 10 ml/kg. Sample size was 5–9 per group. All protocols for handling laboratory animals were approved by the Wright-Patterson Institutional Animal Care and Use Committee (IACUC) and meet appropriate Federal guidelines. Fisher 344 rats (~ 250 g) were obtained from Charles Rivers Laboratory equipped with jugular vein catheters, and allowed to acclimate for 7 days prior to the start of the study. Animals were then housed individually in metabolism cages and given ad libitum access to food (Purina Certified Rat Chow # 5002) and water. The housing environment was maintained on a 12 h light–dark cycle at 25°C, and all animals were examined by Vivarium personnel twice daily to ensure their health and well-being. Urine was collected into containers chilled on dry ice and containing 1 ml of 1% sodium azide. All urine samples were stored at –40°C prior to analysis by NMR spectroscopy.

Urine samples for NMR analysis were prepared as described by Robertson et al. (Robertson et al. 2000) and modified as follows (Westrick et al., submitted). Samples were thawed at 4°C overnight then allowed to equilibrate to room temperature just prior to NMR sample preparation. A 600 μl aliquot of urine was mixed with 300 μl of a phosphate buffer (0.2 M monosodium phosphate and 0.2 M disodium phosphate, pH 7.4) and allowed to equilibrate for ten minutes. Samples were then centrifuged at 5000 rpm (2300 rcf) for ten minutes to remove any particulate matter and 550 μl of supernatant was transferred to a 5 mm NMR tube. An internal standard consisting of 150 μl of trimethylsilylpropionic (2, 2, 3, 3 d₄) acid (TSP) dissolved in deuterium oxide was added at a final concentration of 2 mM.

Proton NMR spectra were acquired at 25°C on a Varian INOVA operating at 600 MHz. Water suppression was achieved using the first increment of a NOESY pulse sequence, which incorporated saturating irradiation (on-resonance for water) during the relaxation delay (7.0 s total; 2 s with water presaturation) and the mixing time (50 ms total; 42 ms with water irradiation). Data were signal averaged over 64 transients using a 4.0 s acquisition time and interpulse delay of 11.05 s.

NMR spectral data were processed using Varian software and employing exponential multiplication (0.3 Hz line-broadening), Fourier transformation, and baseline flattening (fifth-order polynomial and spline fitting routines). The TSP signal was used as an internal chemical shift reference (set at 0.0 ppm), and the regions

surrounding the residual water signal (≈ 4.8 ppm) and the urea signal (≈ 5.8 ppm) were excluded from the analyses. The vertical shift of the entire spectrum was adjusted such that the mean of the intensities between 11.6 and 10 ppm (a region containing spectral noise) was zero. Then the peak intensities of each spectrum were normalized to a constant sum.

2.2 Algorithm

Dynamic adaptive binning determines the optimal bin configuration of n observed peaks as measured by an objective function. This process is divided into two steps: (1) determining the location of the observed peaks in each spectra and (2) finding the optimal bin boundaries with respect to the objective function. The identification of the observed peaks in each spectrum is accomplished by identifying local maxima after smoothing via a wavelet transform (Alsberg et al. 1997; Cancino-De-Greiff et al. 2002; Kaczmarek et al. 2004; Perrin et al. 2001; Shao et al. 2003). After the observed peaks of each spectrum have been determined, the algorithm determines the optimal bin configuration using a dynamic programming strategy to efficiently find the best solution. These bin boundaries can then be used to quantify additional spectra.

2.2.1 Optimizing bin boundaries using dynamic programming

The complexity of identifying significantly responding metabolites (i.e., biomarkers) is increased when multiple peaks fall in the same bin. Ideally, each bin should contain a single peak from each spectrum representing the same metabolite. In ^1H NMR spectra, a peak representative of a single type of proton in a molecule (i.e., methine, methyl, etc.) can sometimes be split into a multiplet (i.e., doublet, triplet, etc.) due to J-coupling. Our approach does not attempt to address this issue, but rather, bin boundaries are selected to ideally contain only a single peak. As an optional post binning step, the user can interactively modify bin boundaries in an effort to combine signals that are identified as a J-coupled multiplet. The degree to which a bin approaches this ideal is approximated by counting the number of observed peaks within its boundaries for each spectrum. This can be quantified by a bin heuristic objection function (*BHOF*) that is calculated as follows:

$$\text{BHOF}(\alpha, \omega) = \sum_s |1 - N_s|, \quad (1)$$

where N_s is the number of observed peaks in spectrum s for the region defined by the bin boundaries, $[\alpha, \omega]$. A *BHOF* value of 0 indicates that for the bin $[\alpha, \omega]$ each spectrum has one observed peak.

The bin heuristic objective function measures the fitness of an individual bin. For a set of bins, β , a global heuristic objective function (*GHOF*) is calculated as follows:

$$\text{GHOF} = \sum_{[\alpha, \omega] \in \beta} \text{BHOF}(\alpha, \omega), \quad (2)$$

where β is the set of all bins and $[\alpha, \omega]$ are the boundaries of a bin. Thus, two or more sets of bin boundaries, $\{\beta_1, \dots, \beta_n\}$, can be ranked according to their *GHOF* scores. The *GHOF* score represents the cumulative score of the individual bins. The fitness of an individual bin is measured as the degree to which it conforms to the ideal that a bin contains one observed peak from each spectrum.

The *GHOF* score is a discrete function, where different sets of bin boundaries can yield the same *GHOF* score. To decide between these configurations, various tiebreaking heuristic objective functions (*THOF*) can be developed. One way to distinguish between these configurations is by the number of bins, where $\text{THOF} = \|\beta\|$. The choice to maximize or minimize *THOF* will depend on the preference of the researcher, as increasing the number of bins increases the probability of peaks spanning bin boundaries; however, decreasing the number of bins increases the probability of two or more peaks residing in a single bin. For this manuscript, the *THOF* metric is maximized. If two bin configurations have an equal number of bins, then the configuration that maximizes the margins between adjacent bins is selected, where the margin between two adjacent bins is the minimum distance between their observed peaks. The average margin (*AVGM*) is defined as follows:

$$\text{AVGM} = \frac{1}{\|\beta\|} \sum_{i=2}^{\|\beta\|} \text{margin}(i-1, i), \quad (3)$$

where $\text{margin}(i-1, i)$ is the margin between the previous and the i -th bin. Thus, the best binning solution is found by minimizing *GHOF* using *THOF* as a metric to distinguish between equivalent configurations.

In addition to these two objective functions (*GHOF* and *THOF*), unrealistic bin boundaries are avoided by including two additional parameters: the maximum distance between observed peaks in the same bin (W) and the minimum distance between an observed peak and a boundary (D). The first parameter is designed to avoid undesirably large bins, such as combining the entire spectrum in a single bin. The second parameter controls the minimum desired distance between observed peaks in adjacent bins, thus, preventing the algorithm from splitting peaks deemed too close by the user.

The optimal binning of n observed peaks is found via a dynamic programming strategy that minimizes *GHOF* with *THOF* breaking ties. Specifically, the optimal solution for binning n observed peaks is obtained by incorporating the

n-th observed peak into the optimal solution for binning previous observed peaks (illustrated in Fig. 1). Intuitively, the algorithm incrementally builds the solution by incorporating optimal solutions of subproblems. Proofs verifying optimal substructure and overlapping subproblems can be found in the supplemental materials. Formally, the recurrence relationship for incorporating the *n*-th observed peak is given in Table 1.

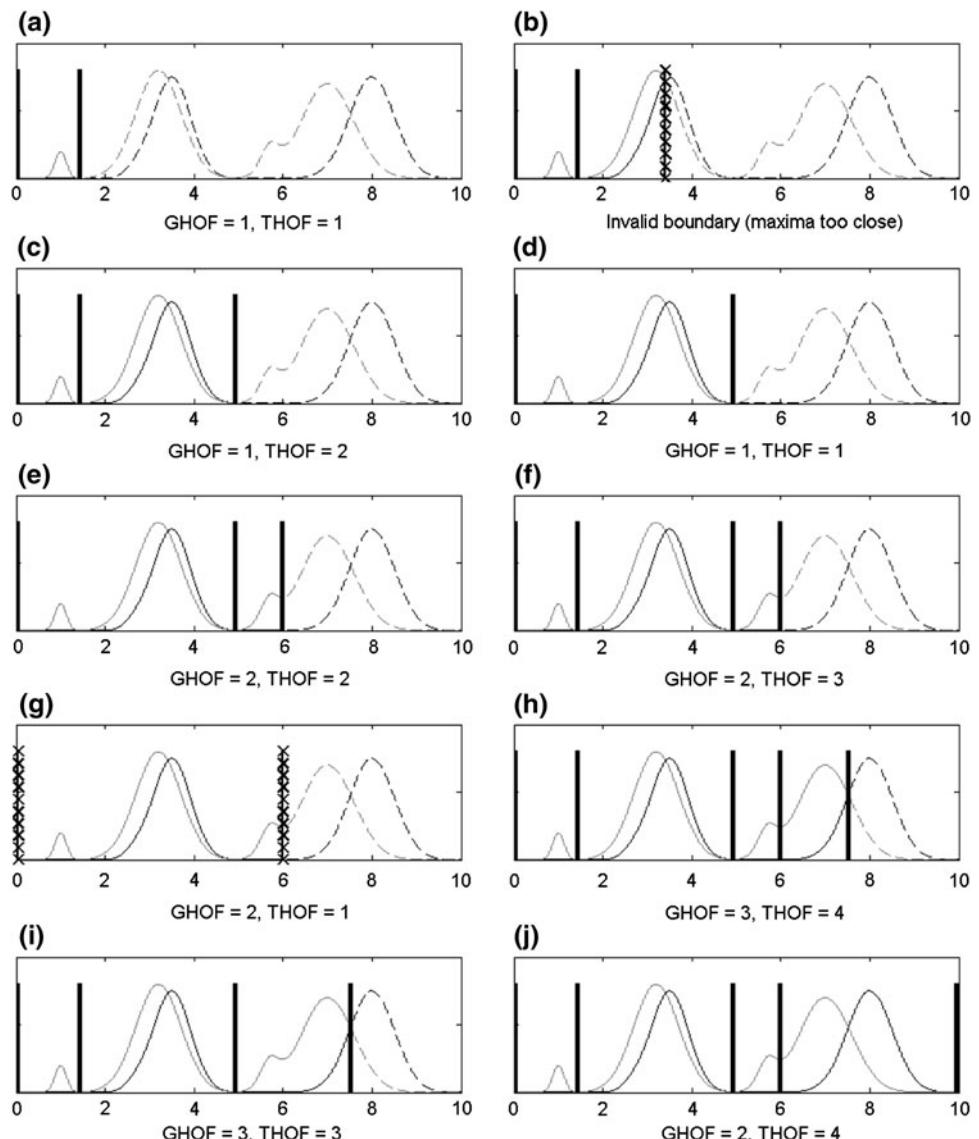
The bin boundaries are defined by the first and last observed peaks in the bin, which are known as the base observed peaks. The left-boundary is defined by finding the location of the minimum intensity of a maximum composite spectrum between the first observed peak in the bin and the previous observed peak. This procedure is repeated with the right-boundary using the last observed peak in the bin and the next observed peak. If this location results in an observed peak to boundary distance below the user-defined

threshold, then the boundary is centered between the adjacent observed peaks. The exceptions include the left-boundary of the first bin and the right-boundary of the last bin, which are set to a distance of half the maximum distance between observed peaks. Note that adjacent observed peaks closer than two times the threshold between observed peaks can be ignored because a valid boundary splitting the observed peaks does not exist.

2.2.2 Selecting the parameters for identifying observed peaks

The procedure for determining the location of the observed peaks begins by smoothing each spectrum using a decimated wavelet transformation (Alsberg et al. 1997; Cancino-De-Greiff et al. 2002; Kaczmarek et al. 2004; Perrin et al. 2001; Shao et al. 2003). A smooth spectrum is created

Fig. 1 Progression of binning six observed peaks from two spectra. The algorithm optimizes the bin boundaries proceeding from left to right, where a dashed line indicates the portion of the spectrum that the algorithm has not considered. (a) The optimal binning of the 1st observed peak. The configuration shown in (b) is unrealistic because the 2nd and 3rd observed peaks are too close to be in separate bins. (c, d) both have identical GHOF scores for binning the first three observed peaks, but (c) would be preferred due to its higher THOF score. (e, f) both show equivalent configurations for binning the first three observed peaks, but (f) is preferred due to its higher THOF score. (g) is invalid due to the distance between the first and last observed peaks. The optimal binning of the first 5 observed peaks is shown in (h), and finally, the optimal binning of the first 6 observed peaks is shown in (j)



by deconstructing each spectrum using a specific wavelet and then zeroing the wavelet coefficients that are designated as noise. There are several options to consider when smoothing via a wavelet transform, including the selection of the wavelet, the threshold selection rule, soft or hard threshold, and whether or not to include multiplicative threshold scaling.

The wavelets selected for evaluation in this study include the commonly used Haar, Daubechies, Symlets, and Coiflets wavelets (Daubechies 1992). Threshold selection rules based on either Stein's unbiased risk (rigsure), a heuristic variant of Stein's unbiased risk (heursure), a universal threshold (sqtwolog), or minimax thresholding (minimax) are evaluated. The value of using hard or soft thresholding is also evaluated, along with the benefit of using multiplicative threshold rescaling. The threshold rescaling techniques evaluated include no rescaling (one), rescaling using a single estimation of level noise based on first level coefficients (sln), and rescaling done using level dependent estimation of noise (mln). Further, baseline variations and incorrectly smoothed regions can result in spurious observed peaks; therefore, only those observed peaks significantly above the noise of the spectrum are retained. The threshold is calculated as n times the standard deviation of a region of noise. All permutations of the aforementioned wavelet parameters and $1 \leq n \leq 6$ are evaluated to determine the optimal wavelet configuration.

The performance of a wavelet smoothing technique is evaluated by comparing the observed peaks to the correct locations of those peaks. The correct locations are determined from the synthetic data set using spectra without noise. The optimal alignment between the computed and correct observed peaks is then calculated to provide the average distance (AD) from the correct observed peaks. The score for matching two observed peaks is equal to the absolute value of the distance between observed peaks, and the penalty for skipping an observed peak is defined as 0.01 ppm. After an optimal alignment is determined, the average distance between matched observed peaks is calculated as follows:

$$AD = \frac{1}{N_{\text{matched}}} \sum_i |cx_i - sx_i|, \quad (4)$$

where N_{matched} is the number of observed peaks matched between the correct and smoothed spectra, and cx_i and sx_i are the locations of the i -th correct and smoothed observed peaks, respectively. In addition, the alignment provides the percentage of missed peaks (PM), and extra (PE) peaks are also computed:

$$PM = \frac{N_{\text{missed}}}{N_{\text{correct}}}, \quad (5)$$

$$PE = \frac{N_{\text{extra}}}{N_{\text{smooth}}}, \quad (6)$$

where N_{correct} is the number of correct observed peaks, N_{missed} is the number of correct observed peaks that are not matched to a smooth observed peaks, N_{extra} is the number of smooth observed peaks not assigned to a correct observed peaks, and N_{smooth} is the number of smooth observed peaks.

2.3 Evaluating and comparing binning algorithms

2.3.1 Metrics applied to synthetic data sets

Synthetic data sets are employed to provide a statistical basis for comparing binning algorithms. For any statistical comparison on synthetic data to be useful, the synthetic data must accurately characterize the salient features of real data. The synthetic data used in this manuscript are based on urine ^1H nuclear magnetic resonance data (Anderson et al. 2009). In total, 40 data sets each with 20 control and 20 treatment spectra were used to compare three binning algorithms: dynamic adaptive binning (DAB), uniform binning, adaptive binning (AB), and adaptive intelligent binning.

For the application considered here—identification of biomarkers of toxicity—the objective of a binning technique is to increase the effectiveness of biomarker identification. The result of such analysis is a set of bins that have been labeled as significantly responding (i.e., responsive). These responsive bins are then examined to determine which metabolites are reflected by each bin. The complexity of this analysis is increased when multiple observed peaks from a single spectrum reside in a single bin. Further, the closer an observed peak is to a boundary the more its effects are distributed across adjacent bins, and the higher the probability that individual peaks will span bin boundaries.

As spectra contribute more than one peak to a bin, the more difficult it becomes to interpret the results. The ability of a binning technique to achieve this ideal is measured by penalizing each extra or missing observed peak in a bin. This metric is called the normalized number of observed peaks per bin (NNP_s). While similar in calculation to the *BHOF* score described in the methods, the NNP_s metric is calculated using clean synthetic spectra from which noise has been removed. Thus, the exact locations of the observed peaks are known when calculating NNP_s . These are not known to the dynamic adaptive binning technique. Further, to fairly compare algorithms, the set of bins included in the NNP_s metric is limited to those bins containing at least one observed peak (i.e., empty bins are excluded). The normalized number of observed peaks per bin is defined as follows:

$$NNP_s(\alpha, \omega) = |1 - N_s|, \quad (7)$$

where N_s is the number of observed peaks in the clean spectrum s for the region defined by bin $[\alpha, \omega]$. A NNP_s value of 0 indicates that for the bin $[\alpha, \omega]$ in spectrum s contains one observed peak.

When calculating the number of observed peaks, the bins are restricted to those containing at least one observed peak.

In addition to measuring the number of observed peaks per bin, the probability of peaks spanning boundaries must be considered when evaluating binning algorithms. The probability of peaks spanning boundaries is approximated by calculating the distance from each observed peak to the nearest boundary (DPB):

$$DPB_i = |\omega_i - p_i|, \quad (8)$$

where p_i is the location of the i -th observed peak, and ω_i is the location of the nearest boundary. Finally, the time complexity of an algorithm is of practical importance. To measure this, the CPU seconds the algorithm spent in user mode is studied for all 40 data sets.

2.3.2 Comparing algorithms on empirical 1H data set

The dynamic adaptive binning method is compared to uniform, adaptive, and adaptive intelligent binning on its ability to analyze a 1H toxicology data set. The motivation of an adaptive binning technique is demonstrated on two sample regions of spectroscopic data. Further, the results after principal component analysis (PCA)—a common unsupervised latent vector visualization technique—are analyzed for each of the aforementioned binning algorithms. The ability of a binning technique to enhance the results of PCA by improving between group separation and within group scatter is illustrated by the PCA scores plots. The parameters for each algorithm are selected from the results of the normalized number of peaks per bin.

3 Results

3.1 Peak identification via wavelet smoothing

All combinations of the wavelet parameters and techniques previously described were evaluated and ranked according to the average of the percentage of peaks missed and the percentage of extra peaks (AVG). The top 10 peak identification configurations are shown in Table S2 of the supplemental information. The most accurate configuration of wavelet parameters, as measured by the average of PE and PM , is wavelet: sym7, thresholding: soft, rescaling:

rigrsure, level: 1, rescaling: one, and number of noise standard deviations: 5. While not significantly different from the other top wavelet configurations, this configuration is assumed for all future analyses. In practice, any of the top configurations would produce similar results.

3.2 Evaluating and comparing binning techniques

Each spectral binning algorithm is analyzed as a function of their tunable parameters. The process of uniform binning is measured as a function of the bin width. For 1H NMR spectra, a standard bin width is 0.04 ppm. For this analysis, the bin width is varied from 0.01 to 0.06 ppm by 0.01 ppm. The performance of adaptive binning (AB) was measured as a function of the level of the wavelet transform, which is varied between 1 and 6 for adaptive binning. For adaptive intelligent binning (AIB), the parameter R is set to 0.15, 0.5, and 0.85. For dynamic adaptive binning (DAB), the maximum bin width is set to 0.04 ppm and the minimum distance from boundary to observed peak is varied from 0 to 0.004 by increments of 0.001. The wavelet parameters for DAB were selected as the best results from Table S1. For a statistical comparison, the algorithms were ranked according to their performance on the four metrics recorded for the synthetic data sets described in Sect. 2.3.1.

A detailed comparison of the four binning algorithms based on the metrics previously described was carried out using several standard statistical tests. First each of the metrics was tested for normality using the Anderson–Darling test, which rejected normality for the number of observed peaks per bin and the distance from observed peak to nearest boundary ($\alpha = 0.05$). The test failed to reject normality for the other two metrics, and thus, they are assumed to come from a normal distribution. Further, the Levene test showed that the variance of the CPU seconds exhibited heterogeneity ($\alpha = 0.05$).

Using Welch’s variance-weighted one-way ANOVA, the mean the CPU seconds per data set was significant using an alpha of 0.05. The Games-Howell multiple comparison test was used to determine significant differences between algorithms. Algorithm and parameter pairs were sorted

Table 1 Recurrence relationship that recursively defines the optimal binning of n maxima

Description	Score
1. n -th maximum is a base for a new bin	$GHOF_n = GHOF_{n-1} + BHOF(\alpha_n, \omega_n)$
2. n -th and $n - i$ maxima are the bases for a new bin if the distance between the two maxima is less than the maximum bin width, where i is an integer greater than 0.	$GHOF_n = GHOF_{n-i-1} + BHOF(\alpha_{n-i}, \omega_n)$

Table 2 The performance of the binning algorithms and their parameters as measured by the mean CPU seconds/data set relative to the fastest algorithm, where 1 normalized CPU second equals 0.31 s on an Intel Core 2 Duo with 4 GB of RAM

Method and parameters	Mean CPU seconds relative to fastest algorithm	Mean rank
AB, 5	1	1
AB, 3	1.05	1.25
AB, 4	1.05	1.39
AB, 6	1.05	1.47
AB, 2	1.12	2.60
AB, 1	1.29	3.39
Uniform, 0.06	1.82	3.98
Uniform, 0.05	2.18	4.61
Uniform, 0.04	2.71	5.22
Uniform, 0.03	3.65	5.84
Uniform, 0.02	5.41	6.45
Uniform, 0.01	10.88	7.06
DAB, 0.004	473.12	8.33
DAB, 0.003	473.43	8.55
DAB, 0.002	474.71	9.39
DAB, 0.001	512.94	10.13
DAB, 0	9406.47	10.75
AIB, 0.15	42841.35	12.04
AIB, 0.5	47335.35	12.07
AIB, 0.85	50460.65	12.09

For a quick comparison, the slowest algorithm (AIB, 0.85) required approximately 4.3 h to complete on average. Algorithm and parameter pairs were sorted according to the multiple comparison tests on the mean ranks. Significantly different results are shown with different adjacent shading

according to their mean rank using the Kruskal–Wallis test with multiple comparisons. These results are summarized in Table 2, where significantly different results are shown with different adjacent shading. To simplify the table, the mean rank was normalized by dividing each rank by the minimum mean rank.

Using the Kruskal–Wallis test (nonparametric one-way ANOVA), the mean ranks of both the distance from observed peak to nearest boundary and the normalized number of observed peaks per bin were significant using an alpha of 0.05. A multiple comparison test on the mean ranks (analogous to the Tukey–Kramer method) was used to determine any significant differences between algorithms. Algorithm and parameter pairs were ranked according to these tests and are summarized in Table 3, where algorithms with differing ranks are significantly different.

3.3 Empirical ^1H NMR data set

The dynamic adaptive binning method and the traditional uniform binning method were applied to a ^1H NMR-based

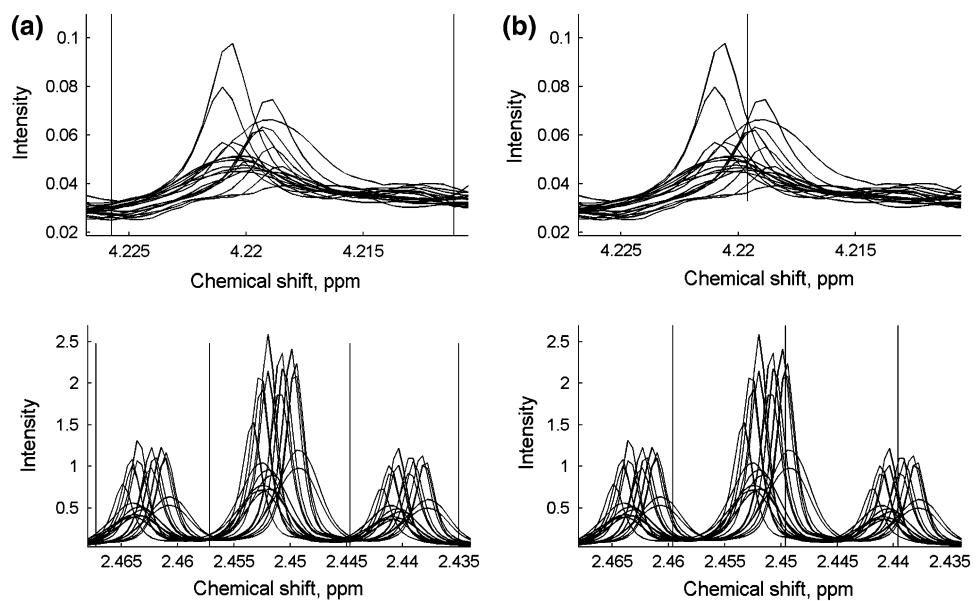
Table 3 The performance of the binning algorithms and their parameters as measured by (a) normalized number of observed peaks per bin, and (b) median distance from observed peak to nearest boundary

(a)		
Method and parameters	Median normalized # of observed peaks/bin	Mean Rank
DAB, 0	0	1
Uniform, 0.01	1	1.96
Uniform, 0.02	1	2.09
DAB, 0.001	1	2.09
AB, 3	1	2.21
AB, 4	1	2.21
AB, 5	1	2.21
AB, 6	1	2.21
AIB, 0.15	0	2.21
AIB, 0.5	1	2.24
AIB, 0.85	1	2.24
AB, 2	1	2.25
AB, 1	1	2.26
DAB, 0.002	1	2.72
Uniform, 0.03	3	2.91
DAB, 0.003	2	2.98
DAB, 0.004	2	3.01
Uniform, 0.04	4	3.39
Uniform, 0.05	5	3.86
Uniform, 0.06	7	4.07

(b)		
Method and parameters	Median dist. from observed peak to nearest boundary	Mean rank
DAB, 0.004	0.48916	4.25
DAB, 0.003	0.16632	3.91
DAB, 0.002	0.05086	3.73
Uniform, 0.06	0.01507	3.19
Uniform, 0.05	0.01249	3.06
Uniform, 0.04	0.01001	2.89
DAB, 0.001	0.00584	2.67
Uniform, 0.03	0.00738	2.66
AIB, 0.85	0.00459	2.39
Uniform, 0.02	0.00499	2.31
AIB, 0.5	0.00250	1.85
Uniform, 0.01	0.00252	1.70
DAB, 0	0.00208	1.58
AIB, 0.15	0.00167	1.52
AB, 6	0.00083	1.02
AB, 5	0.00083	1.02
AB, 4	0.00083	1.02
AB, 2	0.00083	1.02
AB, 3	0.00083	1.01
AB, 1	0.00083	1

Algorithm and parameter pairs were sorted according to the multiple comparison tests on the mean ranks. Significantly different results are shown with different adjacent shading

Fig. 2 Sample regions of ^1H spectroscopic data demonstrating the advantages of dynamic adaptive binning (a) over uniform binning (b)



experiment to monitor rat urinary metabolites. The parameters were selected based on their ability to minimize the normalized number of observed peaks per bin (Table 3b). Figure 2 illustrates the motivation of dynamic binning techniques on two sample regions using a uniform bin width of 0.01 ppm and dynamic adaptive binning parameters $D = 0$ and $W = 0.04$ ppm. The results of principal component analysis on samples from several non-lethal doses of ANIT (20 mg/kg, 50 mg/kg, and 100 mg/kg) are shown in Fig. 3.

4 Discussion

Spectroscopic binning algorithms attempt to enhance the effectiveness of pattern recognition techniques by reducing problem dimensionality with minimal loss of information. One application of a binning algorithm is the determination of biomarkers associated with toxic exposure. The complexity of this analysis increases when multiple peaks fall in the same bin and span bin boundaries, both of which result in a loss of information. Thus, a binning algorithm attempts to minimize the number of dimensions, while maximizing pertinent information and mitigating peak misalignment. The results of this analysis are then analyzed post-hoc to determine the specific metabolites contributing to an individual bin.

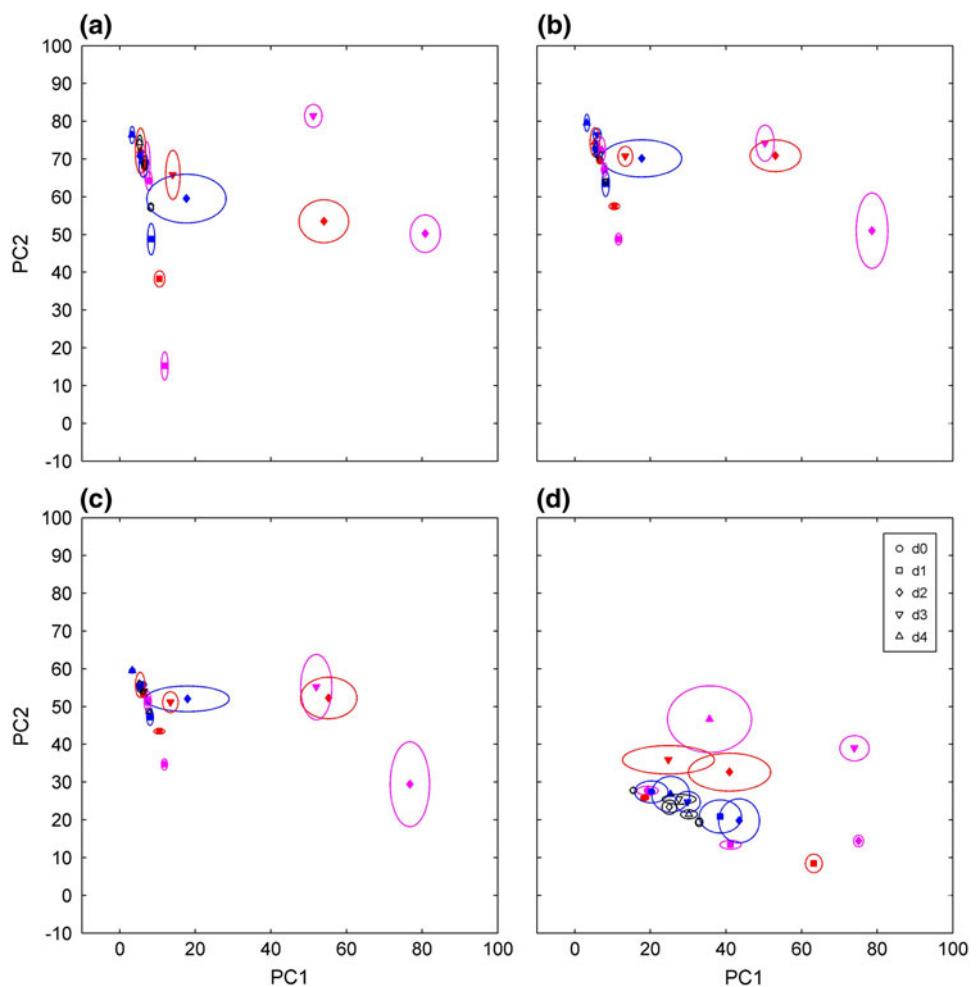
The performance of binning algorithms was measured by the minimum distance from observed peak to boundary (DPB) and the normalized number of observed peaks per bin (NNP) (Table 3). In summary, the dynamic adaptive binning algorithm ($D = 0$) has a significantly better mean rank of normalized number of observed peaks per bin than all other algorithm and parameter pairs. It should also be

noted that while the median normalized number of observed peaks per bin for AIB ($R = 0.15$) is 0, its mean rank is used in the multiple comparison test. When increasing the user-defined parameter, minimum distance from observed peak to boundary ($D = 0, 0.001, 0.002, 0.003$, and 0.004), the probability of peaks spanning bin boundaries decreases, but the normalized number of observed peaks per bin also increases. This increases the complexity of determining the metabolites reflected by each bin. Using a minimum distance from boundary to observed peak of 0.001 ppm balances these two goals.

Secondary to the performance of the algorithms as measured by DPB and NNP , the computational complexity as measured by the CPU seconds spent in user mode is also important. In summary, all methods required significantly less CPU seconds than adaptive intelligent binning ($R = 0.15, 0.5$, and 0.85). In addition, the CPU seconds required by adaptive intelligent binning ($D = 0$), were significantly greater than uniform binning, adaptive binning, and dynamic adaptive binning ($D = 0.001, 0.002, 0.003, 0.004$).

While the advantages of dynamic adaptive binning are quantified using the synthetic spectral data sets, the performance on an experimental ^1H data set is illustrated in Fig. 2. Uniform binning successfully mitigates misalignment when peaks fall in the center of the bin; however, it creates boundaries at fixed intervals, regardless of the environment. This can lead to peaks spanning adjacent bins, as shown in Fig. 2. The probability of a peak spanning bin boundaries decreases as the bin width increases; however, this also increases the probability of multiple peaks residing in a single bin. In general, uniform binning lacks the flexibility to deal with the complexities of a ^1H NMR spectrum.

Fig. 3 Principal component scores (means and standard error) after dynamic adaptive binning (a), uniform binning (b), adaptive binning (c), and adaptive uniform binning (d) for several non-lethal ANIT doses (control: black; 20 mg/kg ANIT: blue; 50 mg/kg ANIT: red; and 100 mg/kg ANIT: magenta). Each dose is measured as a function of time (e.g., d2 is 2 days post-dose). See text for details



The ability of each binning technique to enhance subsequent pattern recognition techniques by improving within and between group scatter is demonstrated by analyzing the PCA results on the ^1H toxicology data set (Fig. 3). This scores plot shows that dynamic adaptive binning provides equivalent or better separation measured by the Euclidian distance between means for each ANIT dose (20 mg/kg, 50 mg/kg, and 100 mg/kg) throughout the time course (day-1, day-2, day-3, and day-4). A specific example of this improvement is the increase in separation between 100 mg/kg ANIT samples at day-3 (d3) and 50 mg/kg ANIT samples at day-2 (d2). The separation between these two groups is 28, 4, 4, and 33 for dynamic adaptive binning, uniform binning, adaptive binning, and adaptive intelligent binning, respectively. Examining this separation shows the advantages of dynamic adaptive binning and adaptive intelligent binning versus uniform and adaptive binning. Further, by examining the separation between the 50 mg/kg ANIT samples at day-2 and day-3 for the binning algorithms shows that dynamic adaptive binning, uniform binning, and adaptive binning provide a twofold increase in group separation versus adaptive

intelligent binning. The exact distances are 42, 39, 42, and 17, respectively.

In comparison to dynamic adaptive binning, adaptive binning and adaptive intelligent binning have fewer user defined parameters. In addition, these algorithms avoid the problem of determining the location of observed peaks; however, finding the locations of the observed peaks has several advantages, including the ability for the user to filter the observed peaks of interest (i.e., based on height). Using the observed peaks also provides the user with domain specific parameters, such as minimum distance from observed peak to the nearest boundary. Finally, the inclusion of observed peaks will facilitate the development of more sophisticated objective functions that can improve quantification by identifying multiplets and assisting in further deconvolution. Specifically, the identified peaks may be supplied as input to a targeted approach that removes metabolites identified with high confidence. The updated spectra could then be processed by a binning approach. The software was written in MATLAB and is available for download at <http://birg.cs.wright.edu/panderson/dab.zip>.

References

- Åberg, K. M., Alm, E., & Torgrip, R. J. O. (2009). The correspondence problem for metabolomics datasets. *Analytical and Bioanalytical Chemistry*, 394, 151–162.
- Alsberg, B. K., Woodward, A. M., & Kell, D. B. (1997). An introduction to wavelet transforms for chemometrists: A time-frequency approach. *Chemometrics and Intelligent Laboratory Systems*, 37, 215.
- Anderson, P. E., Raymer, M. L., Kelly, B. J., Reo, N. V., DelRaso, N. J., & Doom, T. E. (2009) Nuclear magnetic resonance synthetic validation sets. Available from: http://birg.cs.wright.edu/nmr_synthetic_data_sets.
- Anderson, P. E., Reo, N. V., DelRaso, N. J., Doom, T. E., & Raymer, M. L. (2008). Gaussian binning: A new kernel-based method for processing NMR spectroscopic data for metabolomics. *Metabolomics*, 4, 261–272.
- Beckwith-Hall, B. M., Holmes, E., Lindon, J. C., Gounaris, J., Vickers, A., Shapiro, M., et al. (2002). NMR-based metabolomic studies on the biochemical effects of commonly used drug carrier vehicles in the rat. *Chemical Research in Toxicology*, 15, 1136.
- Beckwith-Hall, B. M., Nicholson, J. K., Nicholls, A. W., Foxall, P. J., Lindon, J. C., Connor, S. C., et al. (1998). Nuclear magnetic resonance spectroscopic and principal components analysis investigations into biochemical effects of three model hepatotoxins. *Chemical Research in Toxicology*, 11, 260.
- Brekke, T., Kvalheim, O. M., & Sletten, E. (1989). Prediction of physical properties of hydrocarbon mixtures by partial-least-squares calibration of carbon-13 nuclear magnetic resonance data. *Analytica Chimica Acta*, 223, 123–134.
- Brown, T. R., & Stoyanova, R. (1996). NMR spectral quantitation by principal-component analysis II.—determination of frequency and phase shifts. *Journal of Magnetic Resonance. Series B*, 112, 32–43.
- Cancino-De-Greiff, H. F., Ramos-Garcia, R., & Lorenzo-Ginori, J. V. (2002). Signal de-noising in magnetic resonance spectroscopy using wavelet transforms. *Concepts in Magnetic Resonance*, 14, 388–401.
- Cloarec, O., Dumas, M. E., Craig, A., Barton, R. H., Trygg, J., Hudson, J., et al. (2005). Statistical total correlation spectroscopy: An exploratory approach for latent biomarker identification from metabolic 1H NMR data sets. *Analytical Chemistry*, 77, 1282.
- Connor, S. C., Gray, R. A., Hodson, M. P., Clayton, N. M., Haselden, J. N., Chessel, I. P., et al. (2007). An NMR-based metabolic profiling study of inflammatory pain using the rat FCA model. *Metabolomics*, 3, 29–39.
- Crockford, D. J., Keun, H. C., Smith, L. M., Holmes, E., & Nicholson, J. K. (2005). Curve-fitting method for direct quantitation of compounds in complex biological mixtures using 1H NMR: Application in metabolic toxicology studies. *Analytical Chemistry*, 77, 4556–4562.
- Daubechies, I. (1992). *Ten lectures on wavelets*. Society for Industrial and Applied Mathematics (SIAM).
- Davis, R. A., Charlton, A. J., Godward, J., Jones, S. A., Harrison, M., & Wilson, J. C. (2007). Adaptive binning: An improved binning method for metabolomics data using the undecimated wavelet transform. *Chemometrics and Intelligent Laboratory Systems*, 85, 144–154.
- De Meyer, T., Sinnaeve, D., Van Gasse, B., Tsiporkova, E., Rietzschel, E. R., De Buyzere, M. L., et al. (2008). NMR-based characterization of metabolic alterations in hypertension using an adaptive, intelligent binning algorithm. *Analytical Chemistry*, 80, 3783–3790.
- Defernez, M., & Colquhoun, I. J. (2003). Factors affecting the robustness of metabolite fingerprinting using 1H NMR spectra. *Phytochemistry*, 62, 1009–1017.
- Fiehn, O. (2002). Metabolomics—the link between genotypes and phenotypes. *Plant Molecular Biology*, 48, 155–171.
- Forshed, J., Andersson, F. O., & Jacobsson, S. P. (2002). NMR and bayesian regularized neural network regression for impurity determination of 4-aminophenol. *Journal of Pharmaceutical and Biomedical Analysis*, 29, 495–505.
- Forshed, J., Schuppe-Koistinen, I., & Jacobsson, S. P. (2003). Peak alignment of NMR signals by means of a genetic algorithm. *Analytica Chimica Acta*, 487, 189–199.
- Forshed, J., Torgrip, R. J., Åberg, K. M., Karlberg, B., Lindberg, J., & Jacobsson, S. P. (2005). A comparison of methods for alignment of NMR peaks in the context of cluster analysis. *Journal of Pharmaceutical and Biomedical Analysis*, 38, 824.
- Gartland, K. P., Samins, S. M., Nicholson, J. K., Sweatman, B. C., Beddell, C. R., & Lindon, J. C. (1990). Pattern recognition analysis of high resolution 1H NMR spectra of urine. A nonlinear mapping approach to the classification of toxicological data. *NMR in Biomedicine*, 3, 166.
- Griffin, J. L., Williams, H. J., Sang, E., & Nicholson, J. K. (2001). Abnormal lipid profile of dystrophic cardiac tissue as demonstrated by one- and two-dimensional magic-angle spinning (1)H NMR spectroscopy. *Official Journal of the Society of Magnetic Resonance in Medicine*, 46, 249.
- Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24, 417–441.
- Jolliffe, I. T. (1986). *Principal component analysis*. New York: Springer-Verlag.
- Kaczmarek, K., Walczak, B., de Jong, S., & Vandeginste, B. G. (2004). Preprocessing of two-dimensional gel electrophoresis images. *Proteomics*, 4, 2377.
- Lindon, J. C., Holmes, E., & Nicholson, J. K. (2001). Pattern recognition methods and applications in biomedical magnetic resonance. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 39, 1.
- Martens, H., & Naes, T. (1989). *Multivariate calibration*. London: Wiley.
- Nicholson, J. K., Lindon, J. C., & Holmes, E. (1999). Metabolomics: Understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica*, 29, 1181.
- Nicholson, J. K., & Wilson, I. D. (1989). High resolution proton magnetic resonance spectroscopy of biological fluids. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 21, 444–501.
- Perrin, C., Walczak, B., & Massart, D. L. (2001). The use of wavelets for signal denoising in capillary electrophoresis. *Analytical Chemistry*, 73, 4903–4917.
- Reo, N. V. (2002). NMR-based metabolomics. *Drug and Chemical Toxicology*, 25, 375–382.
- Robertson, D. G., Reily, M. D., Sigler, R. E., Wells, D. F., Paterson, D. A., & Braden, T. K. (2000). Metabolomics: Evaluation of nuclear magnetic resonance (NMR) and pattern recognition technology for rapid *in vivo* screening of liver and kidney toxicants. *Toxicological Sciences*, 57, 326–337.
- Robosky, L. C., Robertson, D. G., Baker, J. D., Rane, S., & Reily, M. D. (2002). In *vivo* toxicity screening programs using metabolomics. *Combinatorial Chemistry and High Throughput Screening*, 5, 651.
- Schoonen, W. G., Kloks, C. P., Ploemen, J. P., Horbach, G. J., Smit, M. J., Zandberg, P., et al. (2007a). Sensitivity of (1)H NMR analysis of rat urine in relation to toxicometabolomics. Part I: Dose-dependent toxic effects of bromobenzene and paracetamol. *Toxicological Sciences*, 98, 271.

- Schoonen, W. G., Kloks, C. P., Ploemen, J. P., Smit, M. J., Zandberg, P., Horbach, G. J., et al. (2007b). Uniform procedure of ^1H NMR analysis of rat urine and toxicometabonomics Part II: Comparison of NMR profiles for classification of hepatotoxicity. *Toxicological Sciences*, 98, 286.
- Shao, X. G., Leung, A. K., & Chau, F. T. (2003). Wavelet: A new trend in chemistry. *Accounts of Chemical Research*, 36, 276.
- Shockcor, J. P., & Holmes, E. (2002). Metabonomic applications in toxicity screening and disease diagnosis. *Current Topics in Medicinal Chemistry*, 2, 35.
- Spraul, M., Neidig, P., Klauck, U., Kessler, P., Holmes, E., Nicholson, J. K., et al. (1994). Automatic reduction of NMR spectroscopic data for statistical and pattern recognition classification of samples. *Journal of Pharmaceutical and Biomedical Analysis*, 12, 1215.
- Stoyanova, R., Nicholls, A. W., Nicholson, J. K., Lindon, J. C., & Brown, T. R. (2004a). Automatic alignment of individual peaks in large high-resolution spectral data sets. *Journal of Magnetic Resonance*, 170, 329–335.
- Stoyanova, R., Nicholson, J. K., Lindon, J. C., & Brown, T. R. (2004b). Sample classification based on Bayesian spectral decomposition of metabonomic NMR data sets. *Analytical Chemistry*, 76, 3666–3674.
- Torgrip, R. J. O., Åring, M., Karlberg, B., & Jacobsson, S. P. (2003). Peak alignment using reduced set mapping. *Journal of Chemometrics*, 17, 573–582.
- Vogels, J. T. W. E., Tas, A. C., van den Berg, F., & van der Greef, J. (1993). A new method for classification of wines based on proton and carbon-13 NMR spectroscopy in combination with pattern recognition techniques. *Chemometrics and Intelligent Laboratory Systems*, 21, 249–258.
- Vogels, J. T. W. E., Tas, A. C., Venekamp, J., & van der Greef, J. (1996). Partial linear fit: A new NMR spectroscopy preprocessing tool for pattern recognition applications. *Journal of Chemometrics*, 10, 425–438.
- Wang, Y., Holmes, E., Nicholson, J. K., Cloarec, O., Chollet, J., Tanner, M., et al. (2004). Metabonomic investigations in mice infected with *Schistosoma mansoni*: An approach for biomarker identification. *Proceedings of the National Academy of Sciences*, 101, 12676–12681.
- Weljie, A. M., Newton, J., Mercier, P., Carlson, E., & Slupsky, C. M. (2006). Targeted profiling: Quantitative analysis of ^1H NMR metabolomics data. *Analytical Chemistry*, 78, 4430–4442.
- Whitehead, T. L., Monzavi-Karbassi, B., & Kieber-Emmons, T. (2005). ^1H -NMR metabonomics analysis of sera differentiates between mammary tumor-bearing mice and healthy controls. *Metabolomics*, 1, 269–278.
- Wold, H. (1966). *Estimation of principal components and related models by iterative least squares* (1st ed.). New York: Academic Press.
- Westrick, M. P., DelRaso, N. J., Raymer, M. L., Anderson, P. E., Mahle, D. A., Neuforth, A. E., et al. (Submitted) Dose and time response metabonomic analyses of α -naphthylisothiocyanate toxicity in the rat. *Chemical Research and Toxicology*.
- Zhao, Q., Stoyanova, R., Du, S., Sajda, P., & Brown, T. R. (2006). HiRes: A tool for comprehensive assessment and interpretation of metabolomic data. *Bioinformatics*, 22, 2562–2564.